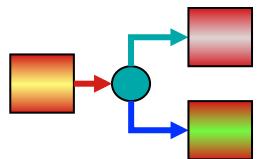


---

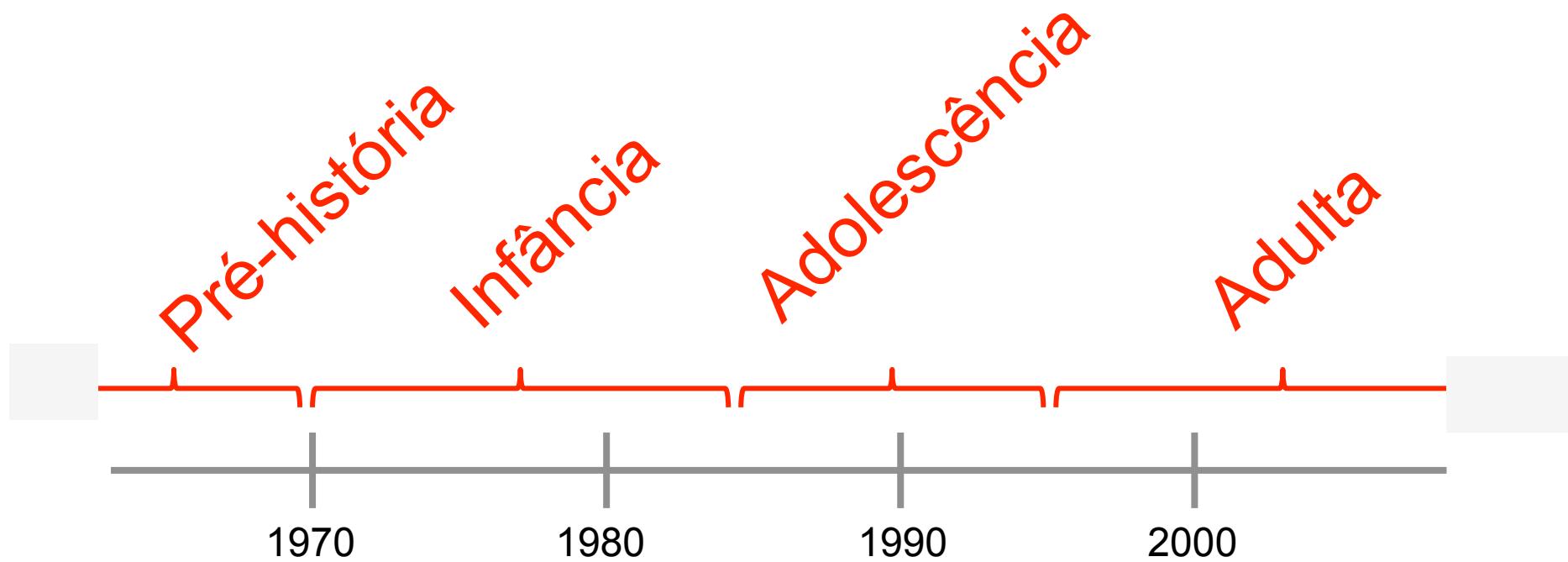
# *Bancos de Dados: de onde vieram, onde estão e para onde irão?*

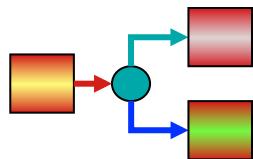
João Eduardo Ferreira ([jef@ime.usp.br](mailto:jef@ime.usp.br))  
Palestra BCC - MAC 0102  
16/05/2019



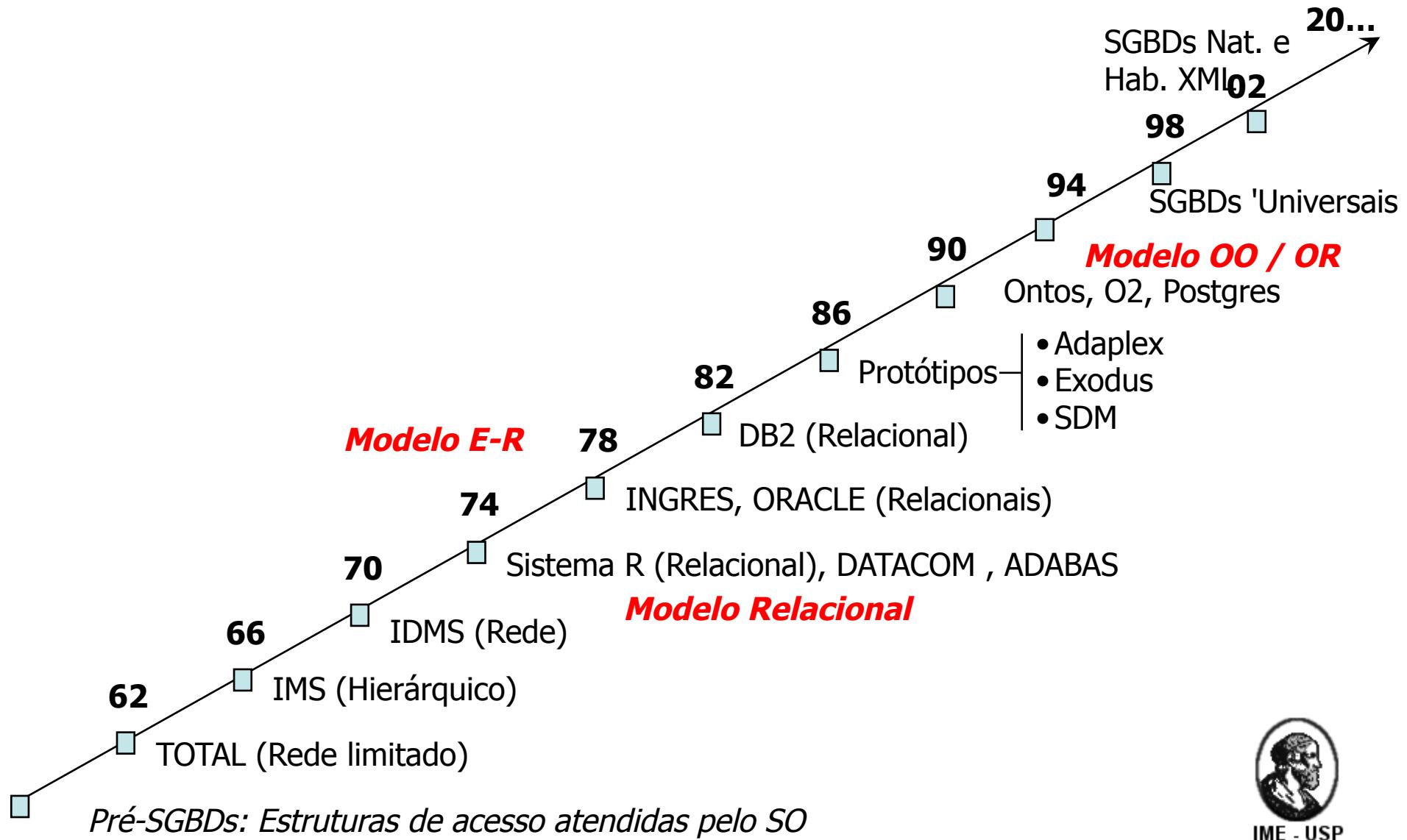


# Linha do tempo de BD

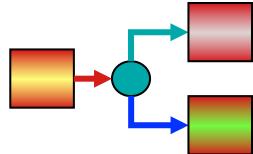




# SGBDs



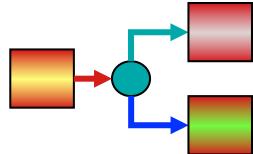
*Pré-SGBDs: Estruturas de acesso atendidas pelo SO*



# Pré-história (antes de 1970)

---

- Limitações de **hardware** →
  - Limitações de **software**
- Baixíssimo nível de abstração
- Sistemas de gerência de arquivos

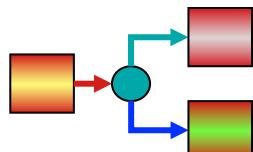


# Como era a área (70-85)

---

- Muitos **recursos** para pesquisa em BD.



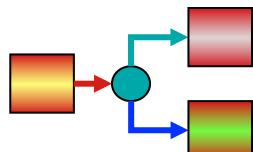


# Como era a área (70-85)

---

- Muitos **recursos** para pesquisa em BD.
- Muitos **grupos** de BD espalhados pelo mundo.





# Como era a área (70-85)

---

- Principal resultado:
  - Modelo de dados **relacional** de Codd (IBM)
  - Técnicas para implementá-lo

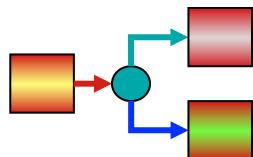
PubID	Publisher	PubAddress
03-4472822	Random House	123 4th Street, New York
04-7733903	Wiley and Sons	45 Lincoln Blvd, Chicago
03-4859223	O'Reilly Press	77 Boston Ave, Cambridge
03-3920886	City Lights Books	99 Market, San Francisco

AuthorID	AuthorName	AuthorBDay
345-28-2938	Haile Selassie	14-Aug-92
392-48-9965	Joe Blow	14-Mar-15
454-22-4012	Sally Hemmings	12-Sept-70
663-59-1254	Hannah Arendt	12-Mar-06

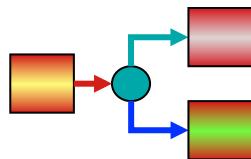
ISBN	AuthorID	PubID	Date	Title
1-34532-482-1	345-28-2938	03-4472822	1990	Cold Fusion for Dummies
1-38482-995-1	392-48-9965	04-7733903	1985	Macrame and Straw Tying
2-35921-499-4	454-22-4012	03-4859223	1952	Fluid Dynamics of Aquaducts
1-38278-293-4	663-59-1254	03-3920886	1967	Beads, Baskets & Revolution



## Principais resultados da pesquisa (70 a 85)

---

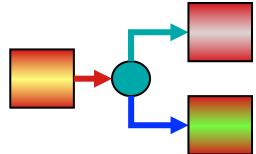
- Teoria de banco de dados
  - resultou na **formalização** do modelo relacional
- SQL
  - linguagem **declarativa** de programação



## Principais resultados da pesquisa (70 a 85)

---

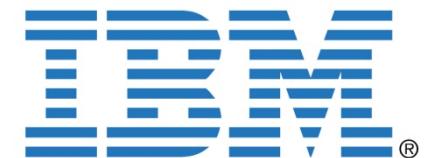
- Otimização de consultas
  - primeira abordagem realista de programação automática
- Gerência de transações
  - algoritmos eficientes para concorrência e reconstrução apoiando grandes bases e grande número de usuários



# Primeiros protótipos e produtos

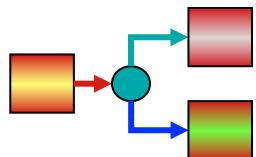
---

- System R (IBM)



- INGRES (Berkeley)





# Primeiros protótipos e produtos

---

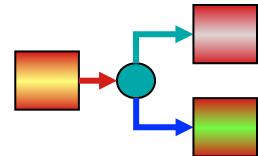
- System R (IBM)



**ORACLE**

- INGRES (Berkeley)

**SYBASE**



# Primeiros protótipos e produtos

---

- System R (IBM)



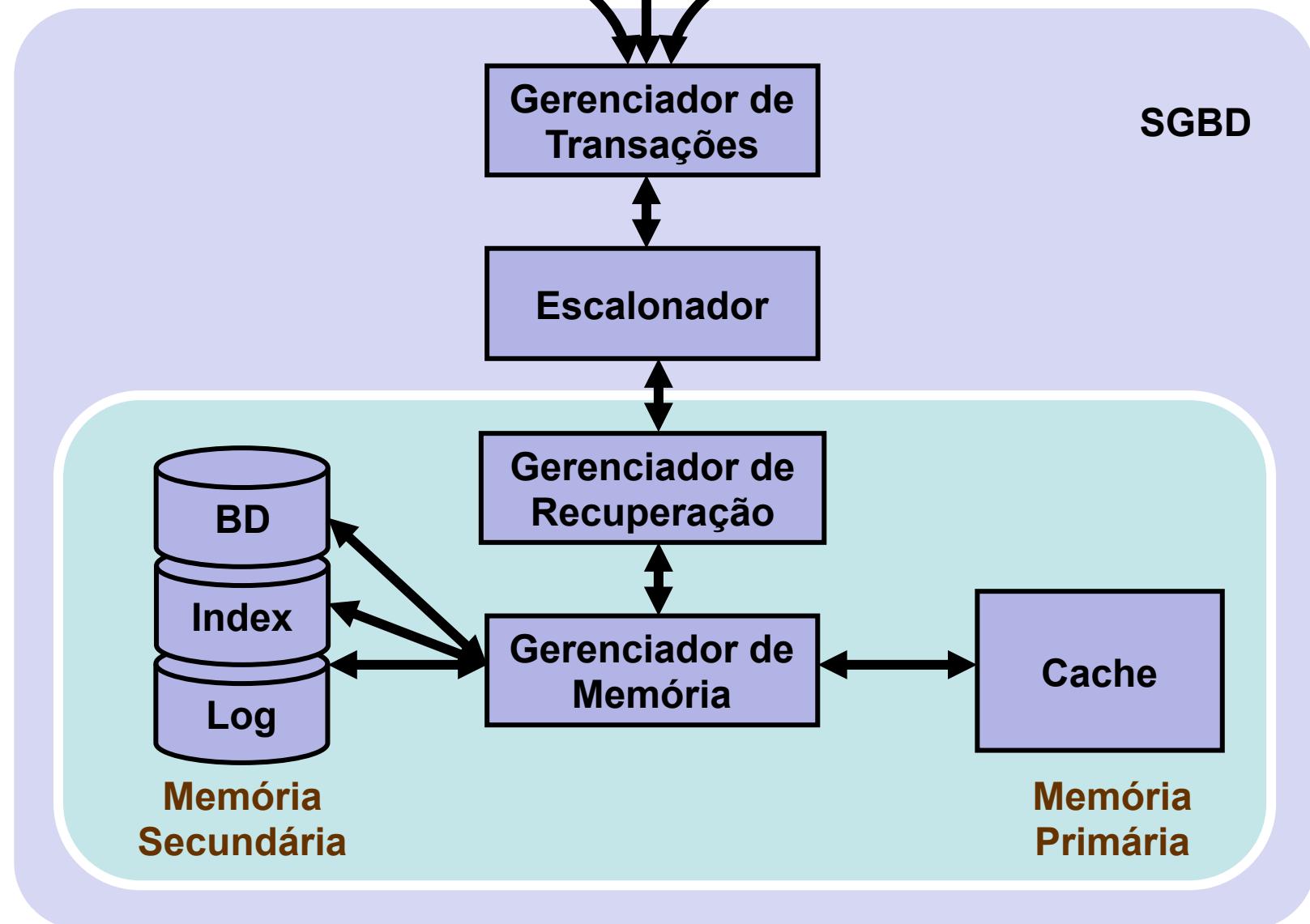
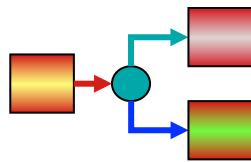
**ORACLE**

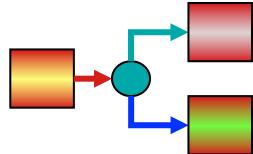
- INGRES (Berkeley)

**Informix**  
SOFTWARE  
The database company™



Microsoft®  
**SQL Server** 2008

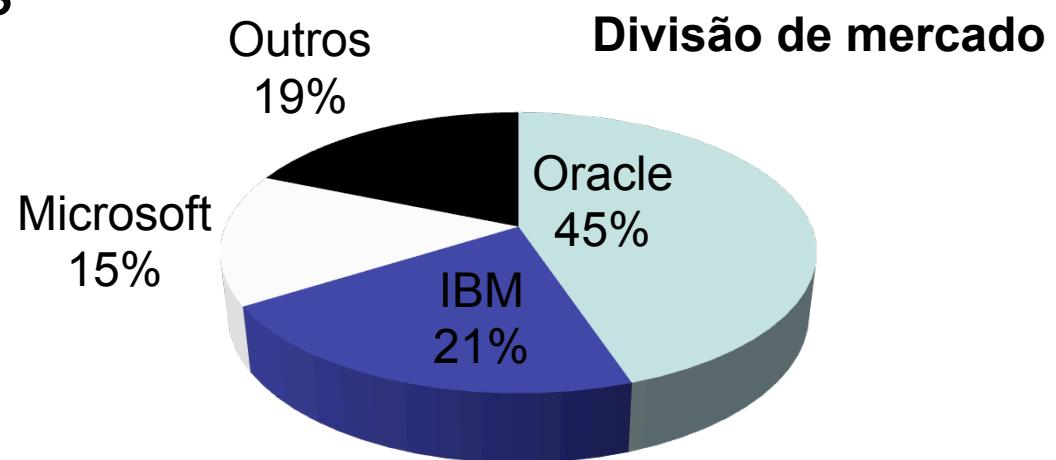


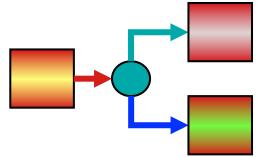


# Indústria de software

---

- Pesquisas na área de BD resultaram em uma importante indústria de software.
- 2017 (IDC):
  - US\$ 60 bilhões

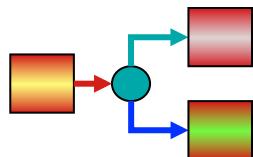




# E no mercado de SGBD?

---

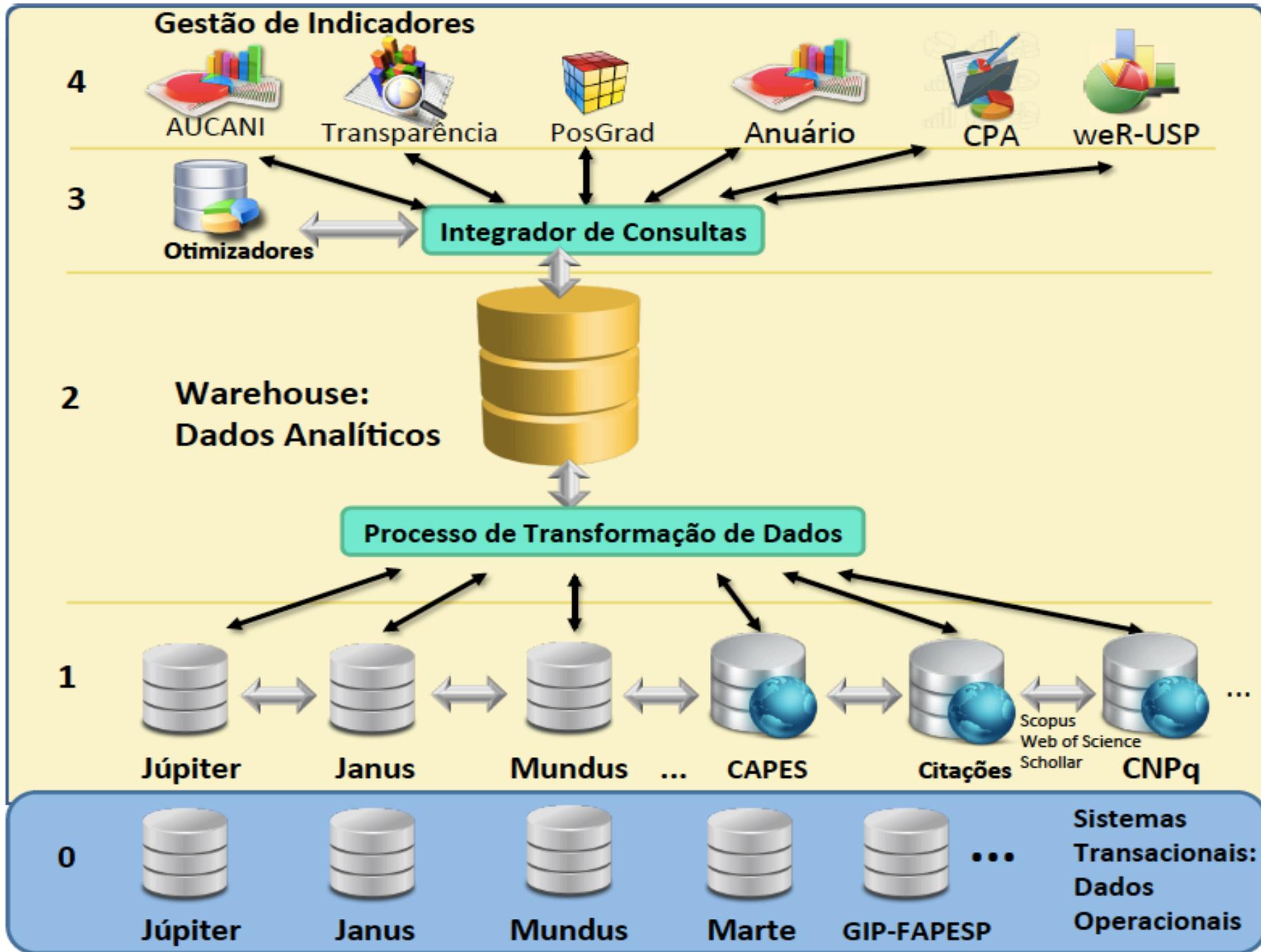
- Mercado de SGBDs **parou de crescer**.
- Grandes usuários de SGBD já o adquiriram: **sistemas legados**.
- Novos usuários:
  - Grande crescimento de bases de dados baseadas em **software livre** (MySQL, PostgreSQL).
- Oracle e Microsoft já distribuem **versões gratuitas**.
- Novo foco das divisões de banco de dados de IBM e Oracle: **Integração de dados**.



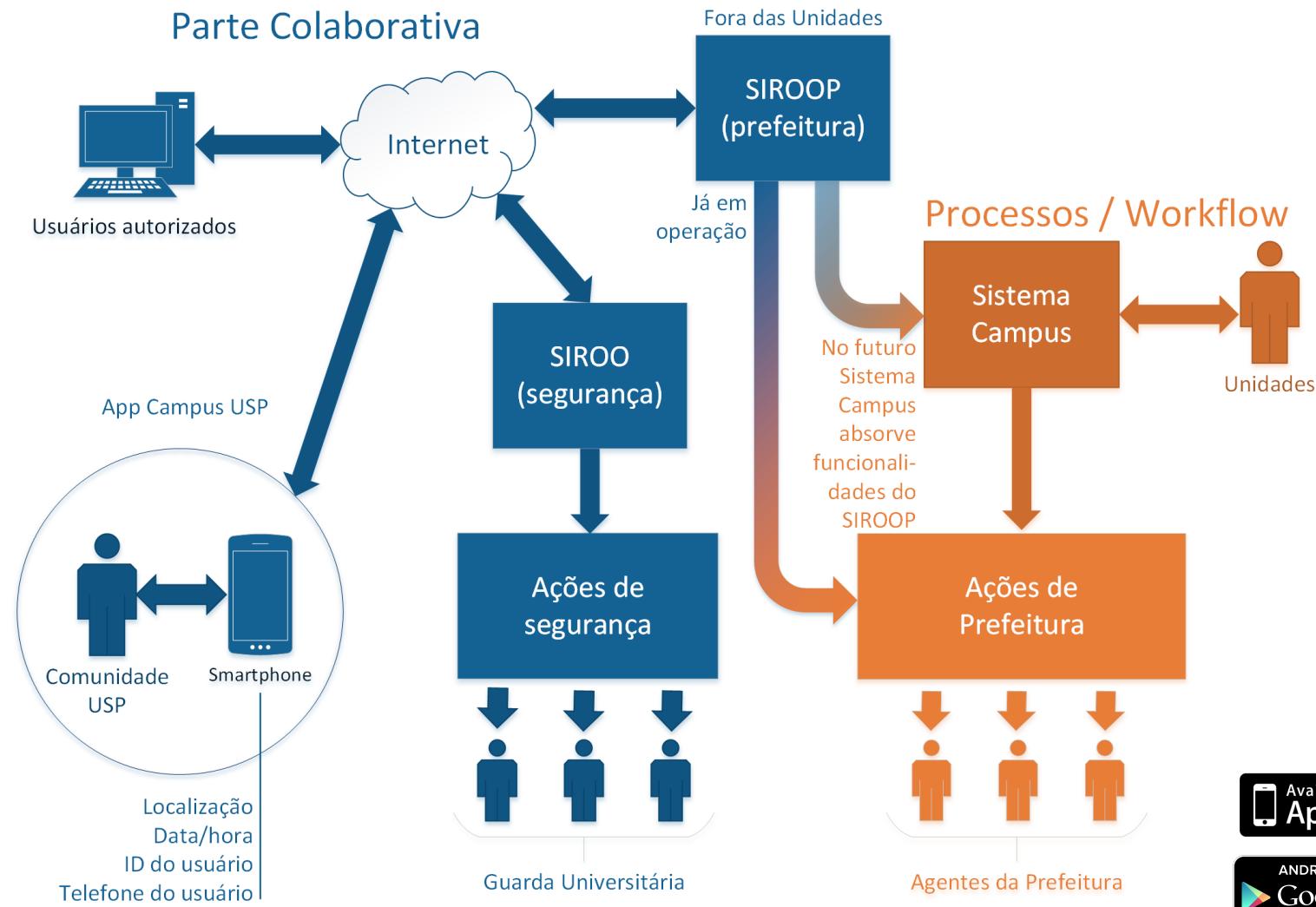
# SGBD onipresente

---

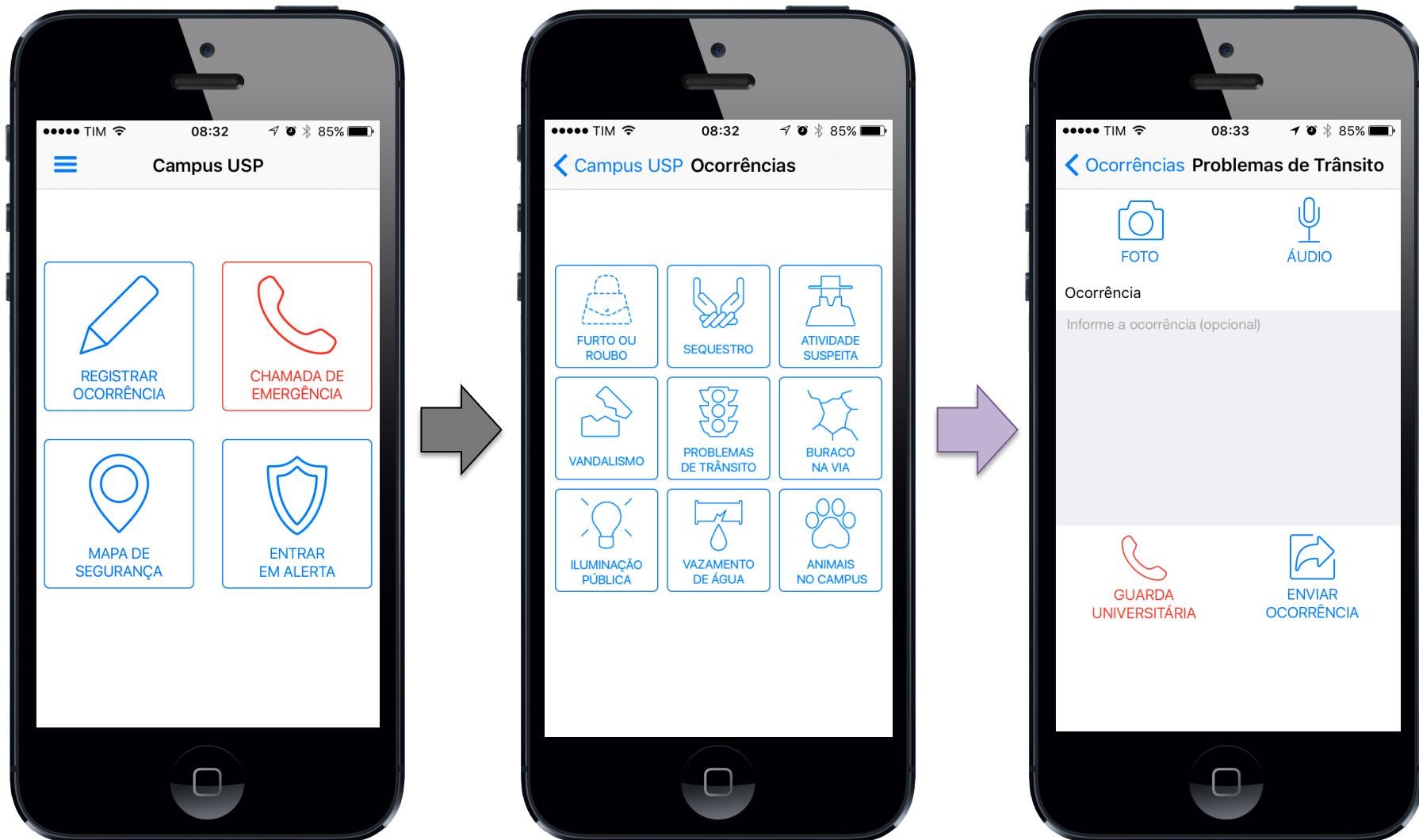
- Originalmente:
  - SGBD era utilizado em **sistemas de informação gerenciais**
- Hoje:
  - SGBD tornou-se **sistema de armazenamento e recuperação em larga escala.**
  - Aparece em **todas** as aplicações



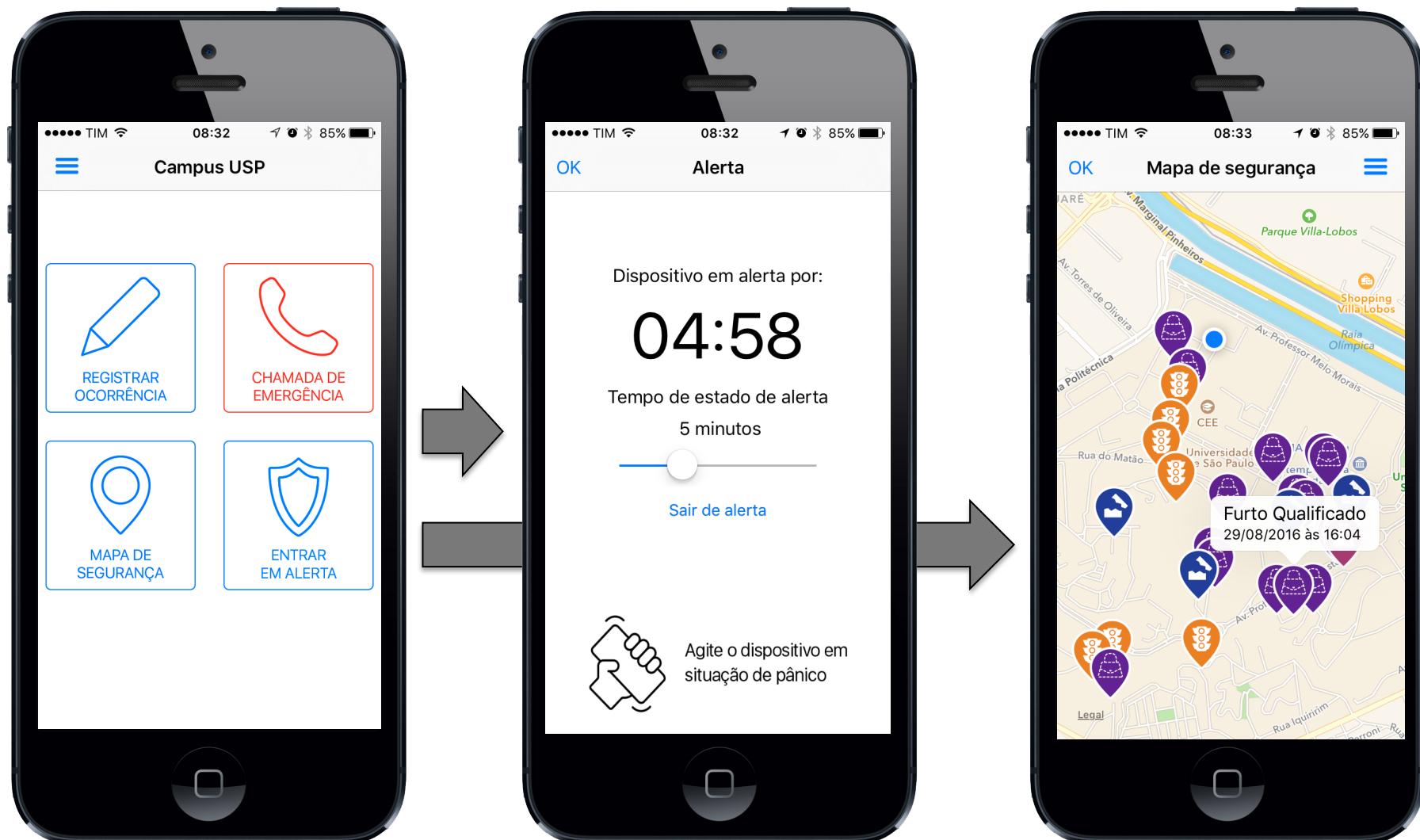
# Sistema Campus de Monitoramento de ocorrências



# Aplicativo móvel para registro e monitoramento de eventos



# Aplicativo móvel para registro e monitoramento de eventos



# Detecção automática de eventos

**P1 ENTRADA FAIXA 1**

EUP5283	Hoje 08:37:32
GEP8015	Hoje 08:37:30
GJM0927	Hoje 08:37:26
FAJ4123	Hoje 08:37:15
FYY7038	Hoje 08:37:09
PUI8642	Hoje 08:37:04
DUG6451	Hoje 08:36:29

S11 - 01 - P1 - ENTRADA

9 FPS (I/P Frames: 20) @ 1 184 KB/s (1.472 Kbps)

Cadastrar nova placa

Hoje 08:37:32 EUP5283

Status da conexão: Monitorando...

**P2 ENTRADA FAIXA 1**

FNN5052	Hoje 08:37:36
FSE7174	Hoje 08:37:32
EAX9565	Hoje 08:37:30
EMW4426	Hoje 08:37:26
QNL1161	Hoje 08:37:24
GA00218	Hoje 08:37:18
GA00210	Hoje 08:37:16

S11 - 13 - P2 - ENTRADA

20 FPS (I/P Frames: 40) @ 1 662 KB/s (5.456 Kbps)

Cadastrar nova placa

Hoje 08:37:36 FNN5052

Status da conexão: Monitorando...

**AV PROF LUCIANO GUALBERTO FX2**

ESW4503	Hoje 08:37:45
GAX7837	Hoje 08:36:55
ACE1083	Hoje 08:36:49
FVJ4050	Hoje 08:36:19
GH09967	Hoje 08:35:51
KDA4880	Hoje 08:35:51
FGZ3091	Hoje 08:35:48

S11 - 17 - Av. Prof. Luciano Gualberto

10 FPS (I/P Frames: 20) @ 1 185 KB/s (1.480 Kbps)

Cadastrar nova placa

Hoje 08:37:45 ESW4503

Status da conexão: Monitorando...

**P3 ENTRADA FAIXA 1**

GJA1849	Hoje 08:37:42
QOK9931	Hoje 08:37:40
QNL0932	Hoje 08:37:36
FPA0865	Hoje 08:37:25
EYD3792	Hoje 08:37:22
FOM6314	Hoje 08:37:20
ERE6152	Hoje 08:37:14

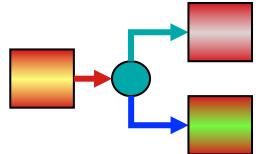
S11 - 25 - P3 - ENTRADA

9 FPS (I/P Frames: 20) @ 1 121 KB/s (1.016 Kbps)

Cadastrar nova placa

Hoje 08:37:42 GJA1849

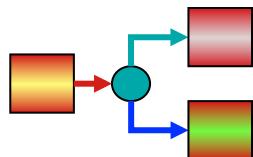
Status da conexão: Monitorando...



# Espaço no mercado

---

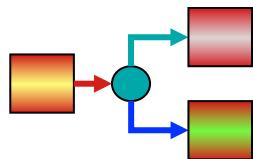
- “Ciência e Engenharia” de dados
  - Aplicações complexas envolvendo grandes massas de dados
  - Web e Nuvem
  - Integração de bancos de dados
  - Visualizações e otimizações de consultas
  - ...



# O que é necessário saber para trabalhar na área

---

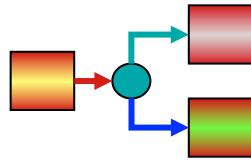
- **Tudo que já era necessário conhecer anteriormente:**
  - Uma boa base de Lógica Matemática
  - Teoria de BD
  - Sistemas operacionais
  - Compiladores
  - Base de matemática discreta
  - Rudimentos de arquitetura de computadores (principalmente, dispositivos de armazenamento)
  - Conhecimentos de BD relacional (linguagens, processamento de consultas, processamento de transações)



# O que é necessário saber para trabalhar na área

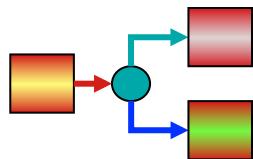
---

- ...., e **muito mais**:
  - Uma base mais ampla de Matemática:
    - álgebra, cálculo, **estatística**...
  - **Paralelismo**;
  - **Recuperação de Informação**;
  - **Aprendizagem de máquina**, . . .



## Espectro do Profissional de BD

- Banco de dados não é mais uma área isolada
- Antes:
  - conhecimentos necessários para atuar na área vinham de alguns **poucos livros-texto**
- Hoje:
  - **amplo espectro** de conhecimentos necessários para atuar com sucesso na área
  - projetos bem sucedidos tendem a incluir pesquisadores de **várias áreas** da computação (IR, ML, IA, paralelismo, ...)



# Bancos de Dados no BCC

---

## Núcleo Básico

- MAC 0350 Desenvolvimento de Sistemas de Computação

## Módulo

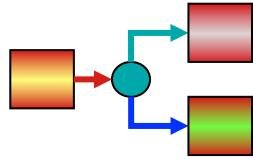
- MAC 0426 Sistemas de Bancos de Dados
- MAC 0439 Laboratório de Bancos de Dados
- MAC 0459 Ciência e Engenharia de Dados

# Conclusões

- Banco de dados deixou de ser uma área de pesquisa **compartimentada** e **isolada** das demais, para tornar-se uma área de pesquisa **multi-disciplinar**.
- **Base matemática** necessária para atuar na área **aumentou** muito.
- Para pesquisar com **sucesso** na área é necessário montar **equipes multi-disciplinares**.

# Conclusões

- A evolução de banco de dados está diretamente vinculada as características dos **dispositivos e tecnologias de hardware para armazenamento** em memória e disco.



## Conclusões

- A área de banco de dados **deixou de tratar apenas o armazenamento de grandes bases de dados em SGBDs relacionais.**
- O foco da pesquisa na área é:
  - tratamento de **grandes massas de dados em SGBDs heterogêneos;**
  - Novos modelos transacionais.